Hellenic Open University
School of Social Sciences
Supply Chain Management

Master Thesis

"Enhancing the monitoring of a vessel economic management through data analytics"

Student/Author:
Mr Karamitros Vassileios

Supervisor professor:
Mr Nikolaos Thomaides

Patra, Winter Semester 2023-2024

Hellenic Open University

School of Social Sciences

Supply Chain Management

Master Thesis

"Enhancing the monitoring of a vessel economic management through data analytics"

Supervising Comittee

Supervisor professor:

Mr Nikolaos Thomaides

Associate Professor

Department of Financial and Management Engineering

University of Aegen

Patra, Winter Semester 2023-2024

I would like to especially thank my family for motivating
to seek my better self via studying and working

Title: Enhancing the monitoring of a vessel economic management through data analytics

Objective: The dissertation's objective is to analyse economic data from the operation of diverse types of ships to investigate at what extent variables influence expenses for maintenance. Ships are ocean-going tanker and bulk carriers that travel all around the globe in ballast or laden condition all year. Data to be used are expenses for purchasing of spare parts, transportation to several ports and relevant custom expenses at these ports. By processing those data, we are hoping to identify the main drivers that influence those expenses.

Data description: Our data consist of records of money spent for three ships of different origin and type. The cost includes money spent mainly for spare parts, freight and stores

Research questions: Our dissertation will address the following (tentative) list of research questions:

- How much are expenses related to type of vessel?
- How much are expenses related to origin of vessel?
- Is any periodicity/seasonality observed in the data? If yes at what time period?
- What is the accuracy of a forecasting model?

All the above research questions will be thoroughly investigated through an array of time series techniques such as regression models, moving average techniques (SMA, WMA, EWMA).

Table of contents

Catalogue of tables and matrixes

Catalogue of diagrams

1. An Exploration of the Historical Antecedents, Categorization, and Design Principles of Cargo Ships

Maritime travelling and trading is as old as human history. One of the oldest pieces of evidence of boat can be found in Norway, Efjordern fjord, which shows "*a life size representation of a boat made from sealskin*" (*Gershon*, 2021) and is almost 11000 years old. One of the first civilizations that made sea voyages was the Egyptian one, who built boats made of papyrus reeds and were propelled either by oar or sails. Several paintings depict Egyptian ships and can be found in ancient vases and murals which are dated around the year 6000 BCE. Egyptian boats did not only sail the Nile River but "*Under Rameses III, Egyptians made a crossing of the Indian Ocean.*"(*Egyptian Ships - Ages of Exploration*, 2014). Established around Mediterranean Sea, Greeks and Phoenician developed ships capable of sea going travelling and, in that way, they dominated sea trading for centuries. Those two civilizations built not only ships for trading but for warfare too, with the Greek triremes to be an iconic ship. For centuries there was no special development in ship building. Material remained the same, wood and the main way of propelling the ships was either sails, utilizing the wind or oars. During medieval times, the Viking longboat appeared in Europe (during the 8^{th} - 11^{th} centuries) and later 15^{th} and 17^{th} centuries the caravels and the galleons, with the last being the ships that played a crucial role in the ocean travels, which needed durable vessels with a bigger capacity. Out of Europe, several types of ship appear like the junk ships in China which evolved from 3^{rd} century BCE and "*They were found — and in lesser numbers, are still found — throughout Southeast Asia and India, but primarily in China*" (Wikipedia contributors, 2023a). The common between above mentioned types is that the main way of travelling was the wind (sails) or human power (oars). During the industrial revolution, the first steamships make their appearance. The first demonstration of a working steamboat model was in 1787 built by John Fitch but the first successful design appeared two decades later by Robert Fulton (*Steamboat*, n.d.). From that moment engines started replacing sails and oars in ships and made then gradually more complicated. Simultaneously ships were capable to carry more goods at greater distances and thus maritime transportation bloomed. Complicated design led to a more sophisticated way of managing a vessel especially when a ship was not owned by its captain but by a company which could be in a distant country.

Ship categories and general design of a cargo ship

Considering the purpose that a vessel is built, ships can be categorized as below (Ship | *Definition, Types, Old, & Facts*, 2023)

1. Warships
2. Service vessels
3. Miscellaneous
4. Industrial
5. Passenger carriers
6. Cargo carriers

Each of the above categories, include sub-categories. Cargo vessels, also known as freighters or merchant ships, form the backbone of global trade by transporting goods across oceans and waterways. These vessels are designed to carry a wide range of cargoes, from bulk commodities like grains and ores to containerized goods and liquid chemicals. Cargo vessels can be categorized into several subtypes based on their design and intended use:

1. General Cargo Ships: General cargo ships are versatile vessels designed to carry a variety of cargo types, including packaged goods, machinery, and vehicles. They typically feature multiple cargo holds and are equipped with cranes or cargo handling gear for loading and unloading operations.

2. Bulk Carriers: Bulk carriers specialize in transporting bulk commodities such as coal, iron ore, and grain in large quantities. These vessels feature expansive cargo holds with minimal internal structure to maximize cargo capacity and facilitate efficient loading and unloading through large hatches.

3. Tankers: Tankers are dedicated vessels designed for transporting liquid cargoes, including crude oil, petroleum products, and chemicals. They are equipped with specialized tanks for storing different types of liquids and require stringent safety measures to prevent environmental contamination.

4. Container Ships: Container ships revolutionized global trade by introducing standardized containers for transporting goods. These vessels feature large deck spaces with rows of

standardized container slots, allowing for efficient loading, stacking, and unloading of containers using specialized cranes and handling equipment.

5. Roll-On/Roll-Off (RO/RO) Ships: RO/RO ships are designed to transport wheeled cargo such as cars, trucks, and trailers, which can be driven on and off the vessel via ramps. These ships are commonly used for transporting vehicles between ports and are equipped with spacious vehicle decks and efficient loading facilities.

Cargo vessels share common design principles aimed at ensuring the safe and efficient transportation of goods across long distances. Key design considerations include:

- Hull Design: Cargo vessels feature sturdy hulls designed to withstand the rigors of ocean voyages and adverse weather conditions. The hull shape varies depending on the vessel type, with bulk carriers featuring wide, flat bottoms for stability and container ships featuring sleek, streamlined designs for speed and efficiency.
- Cargo Holds: Cargo vessels are equipped with spacious cargo holds designed to accommodate various types of cargo. Bulk carriers feature large, open holds with minimal internal structures to maximize cargo capacity, while container ships feature stacked container bays with integrated securing mechanisms to prevent cargo shifting during transit.
- Propulsion Systems: Cargo vessels are powered by a variety of propulsion systems, including diesel engines, steam turbines, and, more recently, LNG-powered engines. Propulsion systems are chosen based on factors such as fuel efficiency, environmental impact, and operational requirements.
- Navigation and Safety Equipment: Cargo vessels are equipped with advanced navigation and safety equipment to ensure safe passage and compliance with international regulations. This includes radar systems, GPS navigation, AIS transponders, and collision avoidance systems, as well as fire fighting and lifesaving equipment to respond to emergencies.
- Crew Accommodation: Cargo vessels provide accommodation facilities for the crew, including cabins, mess halls, recreation areas, and medical facilities. Crew comfort and well-being are essential for ensuring safe and efficient vessel operations during long voyages.

Each of the above has a distinctive design in order to carry safely the cargo for which was designed, but all these share some common design principles.

A modern cargo vessel is a complex engineering project, designed in every detail to efficiently transport goods across long distances while ensuring the safety and well-being of its crew. The vessel is divided into several distinct areas, each serving a specific purpose and equipped with specialized equipment to fulfil its role in the transportation process.

The accommodation areas of a cargo vessel serve as the living quarters for the crew during their time at sea. These areas are designed to provide comfort and functionality, allowing crew members to rest, eat, and relax during their off-duty hours. Within the accommodation areas, one can find dining rooms where crew members gather to enjoy meals prepared in the galley, which is the ship's kitchen. Recreation rooms provide spaces for leisure activities, while refrigerators and general storage spaces ensure the crew's provisions are kept fresh and readily accessible. Additionally, each crew member is allocated their own room for privacy and personal space. Adjacent to the accommodation areas is the bridge, what may be described as the “brain” of the vessel where navigation and communication activities are conducted. Equipped with advanced navigation systems, radar equipment, and communication devices, the bridge serves as the command centre for navigating the ship safely to its destination.

The engine room of a modern cargo vessel is a fundamental hub of activity for the vessel, housing a many different kinds of machinery and systems which are essential for the vessel's propulsion and operation. At the heart of the engine room exists the main engine, a powerful machine which’s main purposes is to provide enough power in order vessel to navigate safely and in needed speed under all weather conditions.

In addition to the main engine, the engine room is equipped with diesel generators, usually three diesel generators and sometimes there is additionally one shaft generator and all these either together or separately serve as the primary source of electricity for the vessel. Almost all systems that are housed in a ship are powered by these generators, from navigation and communication devices on the bridge to the refrigeration units in the accommodation areas. Without the diesel generators, the vessel would be rendered powerless, unable to operate its critical systems and navigate safely through the waters.

Except of the propulsion and power generation systems in engineer room exist more machinery which are usually described as secondary equipment and those are systems that support the vessel's operation. Fuel oil purification systems ensure that the fuel powering the main engine and generators is free from impurities and contaminants, maximizing engine efficiency and longevity. Transfer and preparation systems handle the complex logistics of fuel management, ensuring a steady supply of fuel to the main engine and generators throughout the voyage under all possible circumstances.

Further to above, the engine room houses pumps and piping systems for handling cargo and ballast water. Cargo pumps, which are part of a tanker vessel but not of a bulk carrier or container carrier, facilitate the loading and unloading of cargo, transferring goods between the ship's holds and the dockside facilities. Ballast water pumps, on the other hand exist in all types of vessel and regulate the vessel's stability and trim by pumping seawater in and out of dedicated ballast tanks, allowing the ship to maintain its balance and manoeuvrability in various sea and loading conditions.

Finally, the cargo spaces of the vessel are dedicated to carrying the goods being transported. These spaces, known as cargo holds for the bulk and container carrier and cargo tanks for the tankers, are designed to accommodate different types of cargo, ranging from bulk commodities like grains and ores to containerized goods and liquid chemicals. The design of the cargo spaces varies depending on the type of cargo being transported, with specialized equipment and securing mechanisms tailored to each specific cargo type. Cargo vessels are equipped with loading and unloading facilities, including cranes, conveyor belts, and hatch covers, to facilitate efficient cargo transfer at all sea conditions and handling operations at ports of call.

2.Operation of a vessel

As mentioned earlier, nowadays a vessel is in most times not owned by its master but by a company that is based far from operation area. Moreover, ships are much more complicated structures compared to the past, loaded with many high-end technological equipment and subject to many different strict safe and environmental operation rules. An operation of vessel assumes sea-going and shore-based personnel. Crew may be part either of engine or deck department. In engine department the highest rank is the Chief Engineer who is responsible for all machinery located on the vessel and he is the one who will decide if what spares are needed and when maintenance of an equipment shall take place. For the deck department the highest rank is the Ship Master or Captain who is responsible for complete operation of the vessel, which included navigation, formalities at port and communication with all participating parties such as company, charters, port authorities etc.

On shore-based personnel, the department are more, and each one specializes in a field and works in co-ordination with the rest. A typical structure of a ship management company includes the below departments

- Chartering
- Operations
- Marine
- Crew
- Technical
- Purchasing department
- IT department
- Accounting department

Chartering department

In the context of the shipping industry, 'chartering' can be described in simple terms as the process or action involving the commercial hiring of vessels, the international transportation of cargoes, and frequently, the strategic pairing of vessels with the suitable cargoes (Plomaritou&Jeropoulos, 2022). Main responsibility for the chartering department is to coordinate and oversee the process of chartering ships to transport cargo. Chartering has the vital role to identify the cargos that vessels can carry efficiently and select the appropriate one which will not cause problem to safe operation of the ship, negotiate, prepare, and sign relevant contracts with charterers and search the market for new profitable opportunities. These opportunities may require some modification for the vessel so appropriate communication with technical department is needed to ensure that ship will be selected for transporting such a cargo. Another crucial role for the chartering department is to compose contracts with brokers and charterers to ensure that the company's maritime interests are protected, operational agreements are clearly defined, and all relevant legal and financial aspects are appropriately addressed.

Operations department

The Operations Department of a ship management company serves as the central nerve center, orchestrating and overseeing the intricate array of maritime activities essential for the smooth operation of the fleet. At its core, this department is responsible for the vigilant planning and execution of vessel operations, ensuring compliance with regulatory requirements and optimization of operational efficiency.

One of the primary responsibilities of the Operations Department is to coordinate voyage planning and route selection, taking into account factors such as weather conditions, fuel consumption, and port availability. Through advanced technological tools and expertise, Operations Managers and Voyage Planners chart the most optimal routes to minimize voyage duration and maximize fuel efficiency, thereby reducing operational costs.

Additionally, the Operations Department maintains close communication with vessel captains to ensure adherence to safety protocols, regulatory requirements, and efficient cargo handling procedures. Port Captains liaise with port authorities to coordinate berthing arrangements, cargo operations, and compliance with local formalities, ensuring that port calls are as fast as possible and efficient cargo handling.

Furthermore, the Operations Department is responsible for monitoring vessel performance in real-time, utilizing advanced technologies and data analytics to optimize operational efficiency and mitigate risks. Through continuous monitoring and analysis and cooperation with Technical department, Operations Managers identify areas for improvement and implement corrective measures, such as hull cleaning, to enhance operational performance and reliability.

Summing up, the Operations Department serves as the main gear of maritime operations within the ship management company, ensuring the seamless coordination of vessel activities, adherence to regulatory requirements, and optimization of operational efficiency.

Safety and Quality Department

The Safety and Quality Department of a ship management company embodies a solid commitment to upholding the highest standards of safety, environmental protection, and operational excellence within the managed fleet. At its core, this department is tasked with developing, implementing, and continually refining robust safety management systems and quality assurance protocols to ensure compliance with international regulations and industry best practices. Safety Managers and Quality Assurance Officers conduct comprehensive audits and inspections onboard vessels, continuously evaluating vessel conditions and crew adherence to safety and environmental rules.

Moreover, the Safety and Quality Department plays a vital role in educating crew members on new regulations and safety protocols, making sure that a culture of safety awareness and continuous improvement is present in all aspects ashore and onboard. Through ongoing training programs and knowledge-sharing initiatives, the department empowers crew members to effectively navigate potential risks and respond to emergency situations with confidence and competence.

In addition to its proactive approach to safety management, the department is actively engaged in crisis management and emergency response preparedness. Safety Managers develop contingency plans and coordinate emergency drills to ensure crew readiness and resilience in the face of unforeseen events such as accidents, environmental incidents, or security threats. Such drills are carried out onboard as a training in a weekly basis and may include scenarios such as fire fighting, recovering and rescuing a seafarer who fell in the sea or even the event of a piracy.

Furthermore, the Safety and Quality Department collaborates closely with regulatory bodies, classification societies, and industry stakeholders to stay in sync of emerging trends, regulations, and best practices in safety and environmental management. This proactive engagement enables the department to anticipate regulatory changes and implement timely measures to ensure compliance and mitigate risks effectively.

In essence, the Safety and Quality Department serves as the department which ensures safety, environmental protection, and operational excellence within the ship management company and the ship.

Crew department

The Crew Department of a ship management company serves as the backbone of personnel management, overseeing various essential aspects related to crew members onboard vessels. Central to its responsibilities is the detailed planning and coordination of crew changes, making ticket arrangements, visa formalities, and local accommodation arrangements until vessels reach port. Additionally, the department is tasked with the management of crew certifications, ensuring compliance with regulatory requirements and facilitating ongoing professional development opportunities for crew members. Through regular communication with vessels, the Crew Department identifies potential opportunities for crew promotion and career advancement, fostering a supportive environment conducive to professional growth.

Moreover, the Crew Department plays a vital role in sourcing and recruiting new qualified crew members as the managed fleet expands, collaborating closely with recruitment agencies and maritime training institutions to identify suitable candidates. Furthermore, the department maintains comprehensive records of crew members' qualifications, experience, and training certifications, ensuring compliance with industry standards and regulatory requirements. This meticulous record-keeping enables efficient crew management and facilitates timely decision-making in crew deployment and rotation.

In addition to its operational responsibilities, the Crew Department is committed to addressing crew welfare concerns and promoting a safe and conducive working environment onboard vessels. This includes facilitating access to medical care, counseling services, and recreational facilities to support crew members' physical and mental well-being during their time at sea. Through its proactive efforts and dedication to crew welfare, the Crew Department contributes significantly to enhancing crew morale, retention, and overall operational efficiency within the managed fleet.

Technical Department

The Technical Department of a ship management company assumes a multifaceted role encompassing various crucial aspects of vessel maintenance, repair, and regulatory compliance. At its core, Technical Managers and Superintendents oversee the planning and execution of maintenance schedules, ensuring the operational integrity and seaworthiness of the fleet. This includes organizing regular vessel inspections and surveys, ranging from hull examinations to machinery test and checks, to meet regulatory and classification society standards. Additionally, they oversee the coordination of drydocks and repairs, liaising with external vendors and contractors to procure spare parts, equipment, and services essential for maintenance activities. In doing so, they exercise budgeting and cost control measures to manage expenses effectively without jeopardizing the operational standards.

Moreover, the Technical Department maintains comprehensive documentation and records for the vessel’s maintenance, repairs, and inspections, ensuring compliance with documentation requirements and facilitating historical analysis. Technical personnel also play an important role in emergency response and crisis management scenarios, developing contingency plans and coordinating emergency repairs to mitigate risks and ensure the safety of both crew and vessel. Furthermore, they follow all technological advancements and innovative solutions in the maritime industry, evaluating new technologies and implementing relevant upgrades or modifications to enhance vessel performance efficiency, and safety but also in order to ensure the commercial availability of the vessel.

In the context of increasing environmental awareness, the Technical Department also prioritizes environmental compliance in coordination with the rest departments, ensuring that vessels stick to environmental regulations and adopt environmentally friendly practices. This includes implementing measures to reduce emissions, prevent pollution, and enhance energy efficiency, aligning with the company's commitment to sustainable maritime operations but following any local or international rules too. Through their proactive efforts and attention to detail, the Technical Department plays a role of paramount importance in achieving the highest standards of safety, reliability, and environmental stewardship within the managed fleet.

Purchasing department

The Purchasing Department of a ship management company plays a critical role in procuring essential supplies, spare parts, and services which are necessary for the maintenance and operation of a vessel and the fleet as a whole. This department is usually divided into two key sub-departments: Supply and Spare Parts, each with distinct roles and responsibilities.

The Supply or Stores Department is tasked with procuring consumable items essential for the well-being of the crew and the smooth functioning of vessel operations. This includes purchasing food, cleaning materials, medical supplies, and other consumables such as tools for example required for daily operations. Supply Officers liaise with suppliers, negotiate contracts, and ensure timely delivery of goods to vessels, thereby ensuring that crew members have access to essential provisions and amenities during their time at sea.

On the other hand, the Spare Parts Department focuses on procuring critical spare parts and equipment necessary for the maintenance and repair of onboard machinery and systems. Spare Parts Officers collaborate with suppliers and manufacturers to source high-quality components, ensuring compatibility with vessel equipment and adherence to industry standards. They maintain comprehensive inventory records, monitor stock levels, and coordinate the timely delivery of spare parts to vessels, minimizing downtime and maximizing operational efficiency. In case of an emergency breakdowns Spare parts department is obliged to find the fastest way to order and deliver onboard the needed parts in order to restore seaworthiness of the vessel.

Moreover, the Purchasing Department as a whole, is responsible for budgeting and cost control related to procurement activities, ensuring that expenditures are within budgetary limits while meeting operational requirements. Procurement Officers analyze expenses, negotiate with suppliers, and implement cost-saving measures to optimize procurement processes and minimize operational costs. Such cost saving measures would be to arrange a local supply of spares instead of transferring parts from a third country in order to avoid freight costs or scheduling a supply in a convenient port where the costs are minimal and the operation of the vessel is not prevented.

In short, the Purchasing Department, with its Supply and Spare Parts sub-departments, plays a vital role in ensuring the availability of needed supplies, consumables, and spare parts which are necessary for the safe and efficient operation of the managed fleet.

Through diligent procurement practices, budget management, and effective supply chain management, the Purchasing Department contributes significantly to the overall reliability, performance, and cost-effectiveness of vessel operations within the ship management company.

IT Department

The IT Department, similarly to other industries, is integral to ensure the efficient and effective operation of various technological systems and processes essential for maritime management either ashore or onboard. This department oversees the development, maintenance, and optimization of IT infrastructure, software solutions, and digital platforms which tailored to the unique needs of maritime operations.

One of the primary responsibilities of the IT Department is to develop and maintain robust IT infrastructure to support critical business functions. This includes managing networks, servers, and communication systems, such as email communication system and satellite communication system, to facilitate constant data exchange and communication between vessels, shore-based offices, and external stakeholders which are for instance port agents or charterers. IT Specialists monitor network performance, troubleshoot technical issues, and implement security measures to safeguard against cyber threats and data breaches, ensuring the confidentiality, integrity, and availability of sensitive information.

Furthermore, the IT Department develops in house or purchase and maintains specialized software solutions and digital platforms to streamline administrative processes, enhance decision-making capabilities, and optimize operational efficiency. This includes implementing enterprise resource planning (ERP) systems, maintenance management software which in short are called PMS, and data analytics tools to automate routine tasks, analyze operational data, and support informed decision-making across various departments.

Additionally to infrastructure and software management, the IT Department plays a crucial role in facilitating digital transformation initiatives and innovation within the organization. IT Specialists evaluate emerging technologies, assess their potential applications in maritime operations, and collaborate with internal stakeholders to implement innovative solutions that drive operational excellence and competitive advantage.

Moreover, the IT Department provides technical support and training to employees to ensure proficiency in using IT systems and software tools effectively. This includes conducting training sessions, for example to newcomers in the form of familiarization, creating user guides, and offering ongoing assistance to address technical queries and issues, empowering employees to leverage technology to optimize their workflow and productivity.

Key takeaways, the IT Department serves as a strategic provider of digital transformation and operational excellence within a company. Through its management of IT infrastructure, development of software solutions, and support for digital innovation, the IT Department plays a role of paramount importance in enhancing efficiency, agility, and competitiveness in maritime management practices.

Accounting department

The Accounting Department of a ship management company assumes a fundamental role in financial management, budgeting, and reporting, ensuring the financial health and regulatory compliance of the organization. This department is responsible for managing financial transactions, maintaining accurate records, and preparing financial statements in accordance with accounting principles and regulatory requirements.

One of the primary responsibilities of the Accounting Department is to manage day-to-day financial transactions, including accounts payable, accounts receivable, and payroll processing. Accountants responsibility is to record financial transactions and ensure timely payment of invoices and salaries, thereby maintaining financial accuracy and liquidity within the organization.

Moreover, the Accounting Department plays a crucial role in budgeting and financial planning, working closely with departmental managers to develop annual budgets, monitor expenditure against budgetary allocations, and identify opportunities for cost-saving initiatives. Financial Analysts analyze financial data, prepare budget forecasts, and provide insights to support strategic decision-making and resource allocation across various departments.

In addition to transaction processing and budget management, the Accounting Department is responsible for financial reporting and compliance with regulatory standards.

Accountants prepare accurate and timely financial statements, including income statements, balance sheets, and cash flow statements, ensuring compliance with accounting principles and regulatory requirements.

Furthermore, the Accounting Department collaborates with internal and external auditors to facilitate financial audits and ensure transparency and accuracy in financial reporting. Auditors review financial records, assess internal controls, and provide recommendations for improvement to strengthen financial governance and mitigate risks.

In conclusion, the Accounting Department serves as the caretaker of financial integrity and accountability within the ship management company. Through its careful management of financial transactions, budgeting, reporting, and compliance, the Accounting Department enables informed decision-making, ensures financial stability, and ensures regulatory adherence, thereby contributing to the long-term success and sustainability of the organization.

3. Thesis description and objectives

The maritime industry plays a crucial role in global trade, transporting vast quantities of goods across oceans. Maintaining the economic efficiency of vessels is vital for both individual shipping companies and the overall competitiveness of maritime transportation. However, accurately monitoring and predicting maintenance expenses can be challenging due to diverse factors such as vessel type, operating conditions, and external forces. The aim of the thesis is to enhance the monitoring of a vessel's economic management through data analytics. This will be achieved by

- Analysis for economic data acquired from different type of vessels (ocean going tanker and bulk carrier) to understand the influence of various variables on expenses needed for the machinery maintenance.
- Identifying the key variables that influence these expenses, including but not limited to:
  - Type of vessel: Does the design, size, or function of the vessel impact maintenance costs?
  - Origin of vessel: Do differences in country of origin and so different machinery makers play a significant role?
- Developing a model to predict future maintenance expenses with increased accuracy. This will enable more informed decision-making regarding maintenance budgets and resource allocation.

The methodology which will be followed, will be a data driven approach in which quantitative analysis techniques will be used such as;

- Data collection and pre-processing: Data on maintenance expenses and vessel characteristics such as origin, age and type will be collected. This data will be cleaned, standardized, and prepared for analysis.
- Exploratory data analysis: Initial analysis will be conducted to understand the distribution of variables, identify potential relationships, and explore trends and patterns.
- Modelling: Regression models will be used to assess the influence of identified variables on maintenance expenses and develop a forecasting model.
- Model evaluation: The accuracy and performance of the forecasting model will be evaluated using appropriate metrics.

It is expected for the research to contribute into the field of maritime economics and management by:

- Providing insights into the basic drivers of maintenance expenses across different vessel types and origins.
- Developing a data-driven forecasting model to improve financial planning and decision-making related to vessel maintenance.
- Offering valuable recommendations for shipping companies to optimize their economic performance through enhanced maintenance management strategies.

The following chapters will delve deeper into the methodology, data analysis, results, and discussion of the findings.

4. Empirical study

The data used in this thesis has been sourced from a Greek-based ship management company, focusing on three distinct ships. Among these vessels are two bulk carriers, which share similarities in size, deadweight tonnage (DWT), and machinery. Notably, these bulk carriers were constructed in two different countries, yet share comparable ages. The third vessel under consideration is a tanker vessel, similarly, aged to the bulk carriers. The dataset encompasses daily records of the monetary value of orders placed over the time of the last 14 years for the mentioned three vessels. Each entry includes not only the order value but also specifies the date on which the order was placed. It's important to note that a single day might witness multiple orders per vessel or even no orders at all.

To streamline the analysis, the daily entries have been consolidated into three-month periods, with the total values of these consolidated entries assigned to corresponding quarters of the year. Each quarter is denoted by a shorthand representation, such as yyyyA, yyyyB, and so forth, signifying quarters within a given year. For instance, the second quarter of 2010 is abbreviated as 2010B.It is pertinent to highlight that in calculating the order value, only the cost of spare parts is considered, inclusive of the associated packaging expenses. The forwarding cost, however, has been excluded from the dataset, ensuring a focused examination of spare part-related expenditures.

Descriptive Statistics

Descriptive statistics are a cornerstone in the field of statistics, offering several vital tools for summarizing, organizing, and interpreting data in a comprehensible manner. At its essence, descriptive statistics provide a means to filter complex datasets into manageable insights, enabling researchers, analysts, and decision-makers to collect valuable information and draw informed conclusions.

Descriptive statistics encompass a variety of techniques and measures designed to characterize the central tendency, dispersion, and distribution of data. These include measures such as mean, median, mode, range, variance, and standard deviation, among others. By employing these statistical tools, analysts can effectively summarize the key features of a dataset, providing a snapshot of its essential characteristics.

One of the primary functions of descriptive statistics is to offer a clear and concise summary of data, allowing for easy interpretation and understanding. This helps to make things simpler and easier for the process of data analysis, making it more accessible to a wider audience. Whether in the fields of scientific research, business analytics, or public policy, descriptive statistics play a crucial role in distilling raw data into actionable insights.

Moreover, descriptive statistics enable researchers to identify patterns and trends within a dataset, shedding light on underlying relationships and correlations. By visualizing data through graphs, charts, and tables, analysts can distinguish meaningful patterns that may not be immediately obvious from the raw data alone. This can inform hypotheses, guide further investigation, and support evidence-based decision-making.

The scope of descriptive statistics is broad, encompassing a wide range of applications across various disciplines. In the sciences, descriptive statistics are used to summarize experimental results, assess the reliability of measurements, and communicate findings to the broader scientific community. In business and economics, descriptive statistics are employed to analyze market trends, track financial performance, and inform strategic decision-making.

Furthermore, descriptive statistics play a crucial role in quality control and process improvement initiatives within industries such as manufacturing and healthcare. By monitoring key performance indicators and analyzing process data, organizations can identify areas for optimization and drive continuous improvement efforts.

In summary, descriptive statistics serve as a fundamental tool for summarizing and interpreting data across diverse fields and applications. By providing a clear and concise summary of data, descriptive statistics enable researchers and decision-makers to extract valuable insights, identify patterns and trends, and make informed decisions.

As mentioned previously, descriptive statistics use several different measures to summarize and describe the main characteristics of a dataset. Some of those measures are going to be listed below along with a brief explanation and description.

Measure of central tendency

Measures of central tendency are a important set of statistical tools in descriptive statistics, offering insights into the typical or central value around which a dataset is distributed.

They provide a summary of the data, aiding in comprehension, comparison, and analysis across various fields. These measures help in understanding the "centre" of the data distribution and are crucial for making inferences and decisions based on data. The primary measures of central tendency include the mean, median, and mode, each with its unique characteristics and applications.

- Mean (Arithmetic Mean): The mean represents the average value of a dataset and is computed by summing all values and dividing by the total number of observations. It is the most commonly used measure of central tendency and is sensitive to extreme values, outliers, making it useful for normally distributed data.
- Median: The median is the middle value of a dataset when arranged in ascending order. Unlike the mean, the median is not influenced by outliers, making it more robust for skewed distributions. So, we may say that is particularly useful when the data contains outliers or when the distribution is not symmetric around a mean value.
- Mode: The mode refers to the most frequently occurring value in a dataset. Unlike the mean and median, the mode can be applied to both numerical and categorical data. It is useful for identifying the most common category or value in a dataset and is not affected by extreme values.

Measures of central tendency play a vital role in summarizing data and understanding its central tendencies. They provide valuable insights into the characteristics of a dataset, such as its typical value and the distribution of observations around that value. Additionally, these measures help in comparing different datasets and identifying patterns or trends within the data. Whether in scientific research, business analytics, or public policy, measures of central tendency are indispensable tools for making sense of data and informing decision-making processes.

Measure of Dispersion

Measures of central dispersion play a crucial role in descriptive statistics, aiding in the assessment of the variability or spread of data. They provide valuable insights into the degree of homogeneity or heterogeneity within a dataset, essentially quantifying how tightly or loosely the data points are clustered around the central tendency. In essence, these measures offer a means to gauge the extent to which the data is spread out or concentrated.

The most common measures of central dispersion include:

- Range: The range is calculated as the difference between the largest and smallest values in the dataset. It serves as a basic indicator of the total spread of the data, yet it is sensitive to outliers, thereby potentially providing a skewed representation of the data's variability.
- Interquartile Range (IQR): The interquartile range is defined as the difference between the 75th and 25th percentiles of the dataset. It offers a more robust measure of spread compared to the range, as it is less influenced by extreme values. The IQR effectively captures the spread of the middle 50% of the data, making it particularly useful for identifying the central dispersion in skewed distributions.
- Standard Deviation: The standard deviation quantifies the average deviation of data points from the mean value. It provides a comprehensive measure of dispersion, taking into account the magnitude of deviations from the mean for each data point. Standard deviation is widely used due to its ability to capture the overall variability of the dataset in a single metric, making it a versatile tool for comparing the spread of different datasets.

These measures, along with others such as variance, quartiles, and visualization tools like diagrams and histograms, collectively contribute to a comprehensive understanding of the distributional characteristics of data. They enable researchers, analysts, and decision-makers to assess the variability within datasets, identify outliers, and make informed interpretations and decisions based on the observed spread of data points. Through their application, measures of central dispersion facilitate rigorous analysis and interpretation of data across a multitude of disciplines and contexts.

Mean and Median calculation

Regarding our dataset, we will start with the calculation of three different means, one for complete dataset, one for each different type of ship, bulk carrier, and tanker and one for each one vessel. Using the excel functions we have the results shown in Table 1;

Table 1. Mean calculation

SUB DATASET	**VALUE**
COMPLETE DATASET	225.29
BULK CARRIERS	175.60
CHINESE BULK CARRIER	228.54
JAPANESE BULK CARRIER	122.66
KOREAN TANKER	324.66

The mean values provide insightful information regarding the average order size across different vessel categories within the complete dataset.

Firstly, the overall mean order size for the complete dataset is 225.29. This value serves as a benchmark against which the mean order sizes of specific vessel types can be compared.

Within the bulk carrier category, the mean order size is slightly lower than the overall dataset mean, at 175.60. This suggests that, on average, bulk carriers have a slightly smaller order size compared to the dataset's overall average.

Among the specific types of bulk carriers, the mean order size for Chinese-built bulk carriers stands out notably higher at 228.54. This indicates that orders for Chinese-built bulk carriers, on average, exceed those of both the overall dataset mean and the mean order size for bulk carriers as a category. Conversely, Japanese-built bulk carriers exhibit a significantly lower mean order size of 122.66, indicating a notable deviation below the dataset's overall mean order size and even below the mean order size for bulk carriers.

In contrast, Korean tankers demonstrate the highest mean order size among all categories, at 324.66. This value substantially surpasses both the overall dataset mean and the mean order size for bulk carriers. The markedly higher mean order size for Korean tankers suggests a substantial preference or demand for this vessel type within the dataset.

In summary, the analysis of mean values highlights distinct patterns and disparities in average order sizes across different vessel categories within the dataset. While Chinese-built bulk carriers exhibit a notably higher mean order size compared to the dataset's overall average, Japanese-built bulk carriers present a considerably lower mean order size.

Meanwhile, Korean tankers stand out with the highest mean order size, indicating a higher expense for machinery maintenance.

Furthermore, a visual representation of the data presented in Table 1 can be observed in Figure 1 below. This graphical depiction facilitates easy identification of the highest mean value, which corresponds to the Chinese-built bulk carrier. This picture shows that there are big differences in the average order sizes for different types of ships. It confirms what we found when we looked at the numbers.

Figure 1. Mean values VS Ship type.

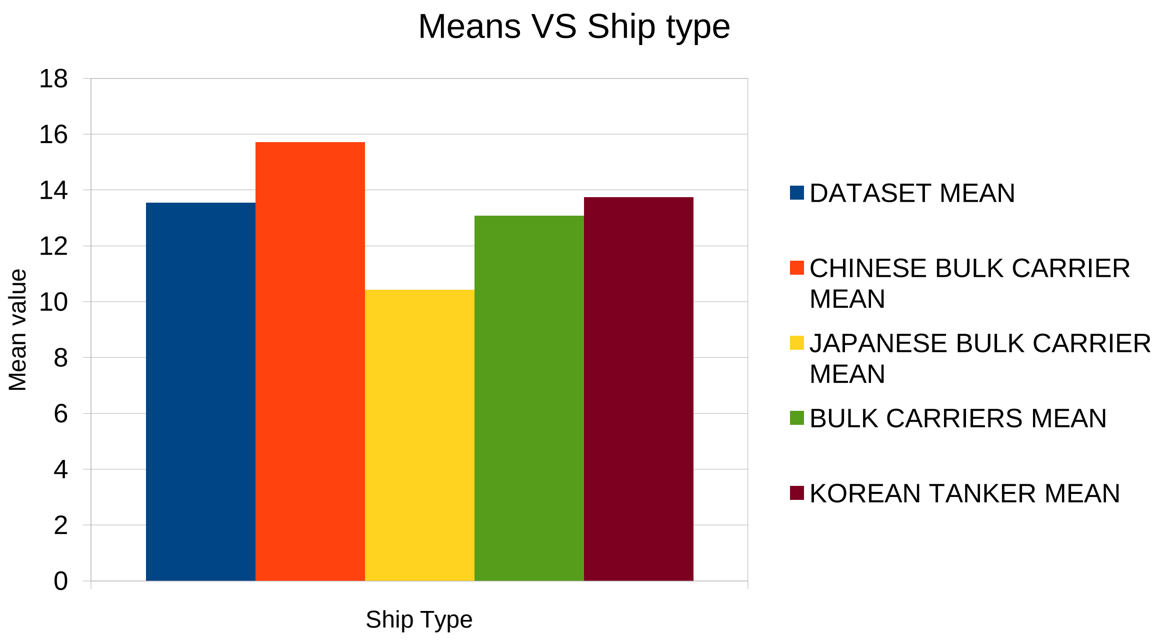


Table 2. Median calculation

SUB DATASET	**VALUE**
COMPLETE DATASET	144.18
CHINESE BULK CARRIER	158.47
JAPANESE BULK CARRIER	68.52
KOREAN TANKER	170.75

The median values presented in table 2, provide valuable insights into the central tendency of order sizes across different vessel categories.

In the complete dataset, the median order size is 144.18. This value represents the middle order size when all orders are arranged in ascending order. It serves as a robust indicator of the typical order size within the dataset.

Among specific vessel categories, Chinese-built bulk carriers exhibit a median order size of 158.47. This indicates that half of the orders for Chinese-built bulk carriers fall below this value, while the other half fall above it. Compared to the median order size of the complete dataset, Chinese-built bulk carriers demonstrate a slightly higher median order size, suggesting a tendency towards larger order sizes for this vessel category.

Conversely, Japanese-built bulk carriers present a notably lower median order size of 68.52. This value is substantially lower than both the median order size of the complete dataset and that of Chinese-built bulk carriers. The comparatively lower median order size for Japanese-built bulk carriers indicates a tendency towards smaller order sizes within this vessel category of bulk carriers but for the whole dataset too.

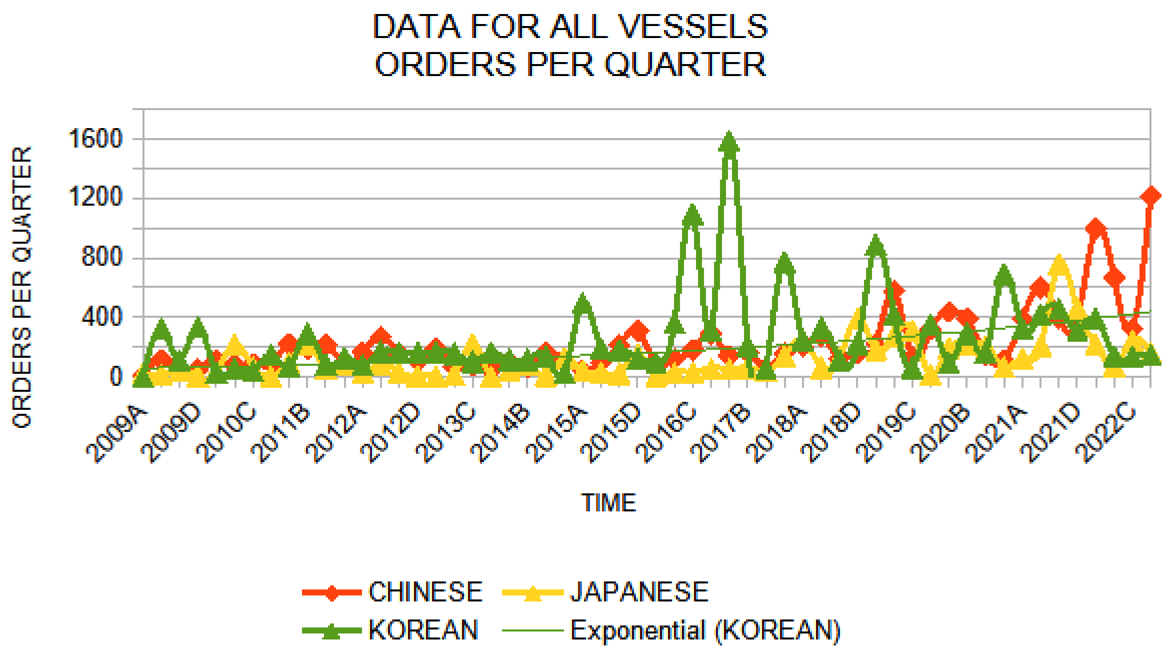
Meanwhile, Korean tankers display a median order size of 170.75, which exceeds the median order size of both the complete dataset and Chinese-built bulk carriers. This suggests a tendency towards larger order sizes for Korean tankers compared to other vessel categories.

In summary, the analysis of median values underscores variations in the central tendency of order sizes across different vessel categories. While Chinese-built bulk carriers exhibit a slightly higher median order size compared to the complete dataset, Japanese-built bulk carriers demonstrate a notably lower median order size. Conversely, Korean tankers present a higher median order size, indicating a need for larger order sizes or more expensive spares.

5. Data plotting

In the following section we will have different diagrams of the data to help understand any connection. We will start with the plot of all data

Figure 2. All data time plot



In Figure 2, all the monetary value of orders placed for three distinct vessels are plotted in a time plot that covers a period of over a fourteen-year period (Q1 2009 - Q4 2022).Overall, we may say that all three ships exhibit a general upward trend in order value over time, suggesting an increase in overall expenses in machinery maintenance. However, fluctuations within the trend point to variations in ordering behaviour. These fluctuations could be influenced by internal factors like planned maintenance cycles or external factors like economic conditions that may lead to increased cost of spare parts due to Covid19 for example.

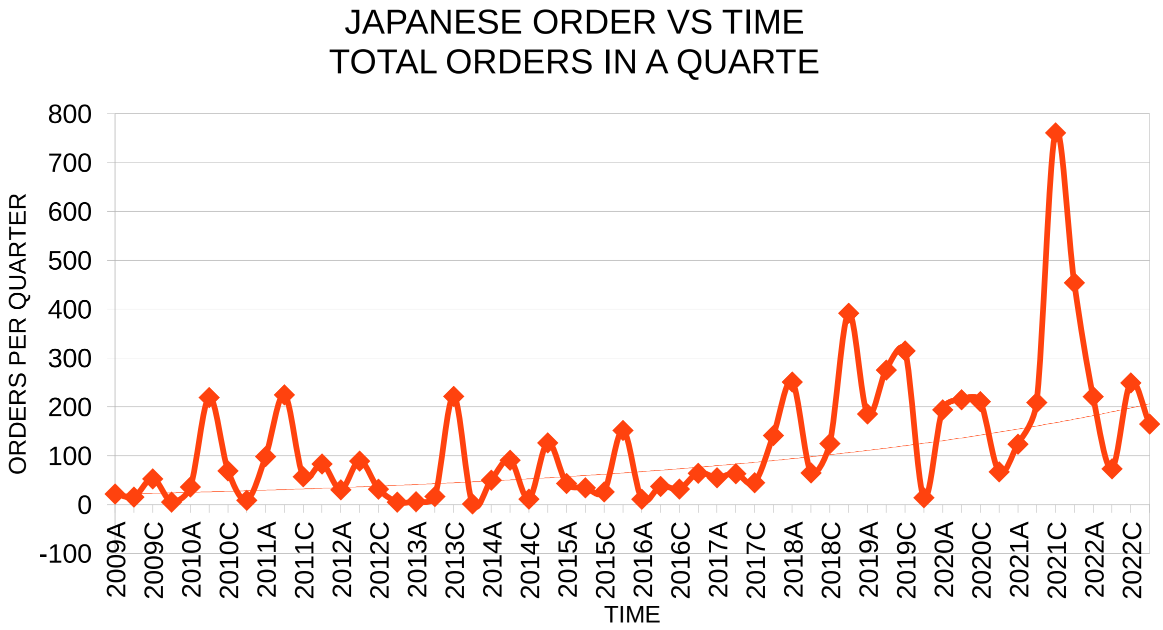
Examining each ship separately, we observe that the Chinese bulk carrier demonstrates a steady rise with minor dips, suggesting consistent maintenance needs that grow as the ship ages. Such a behaviour we may say that is expected. Regarding the Japanese bulk carrier, it is observed the existence of several dips and peaks but of a lesser magnitude. The trend of the expenses for the Japanese vessel is again increasing but not at the same extend as happens in the Chinese one. Last for the Korean tanker vessel, a more dramatic upward trajectory is displayed, particularly in the later years. Generally, we may observe that it compared to the other two ships is more complicated, since until year 2017 the orders' value increases and then starts to be reduced and seems to be stabilized by the end of 2022C.

While the general upward trend suggests increased maintenance needs across all three vessels, individual behaviours differ significantly. This highlights the importance of examining individual trends and considering external factors for a comprehensive understanding of observed patterns.

Figure 3. Chinese bulk carrier orders time plot

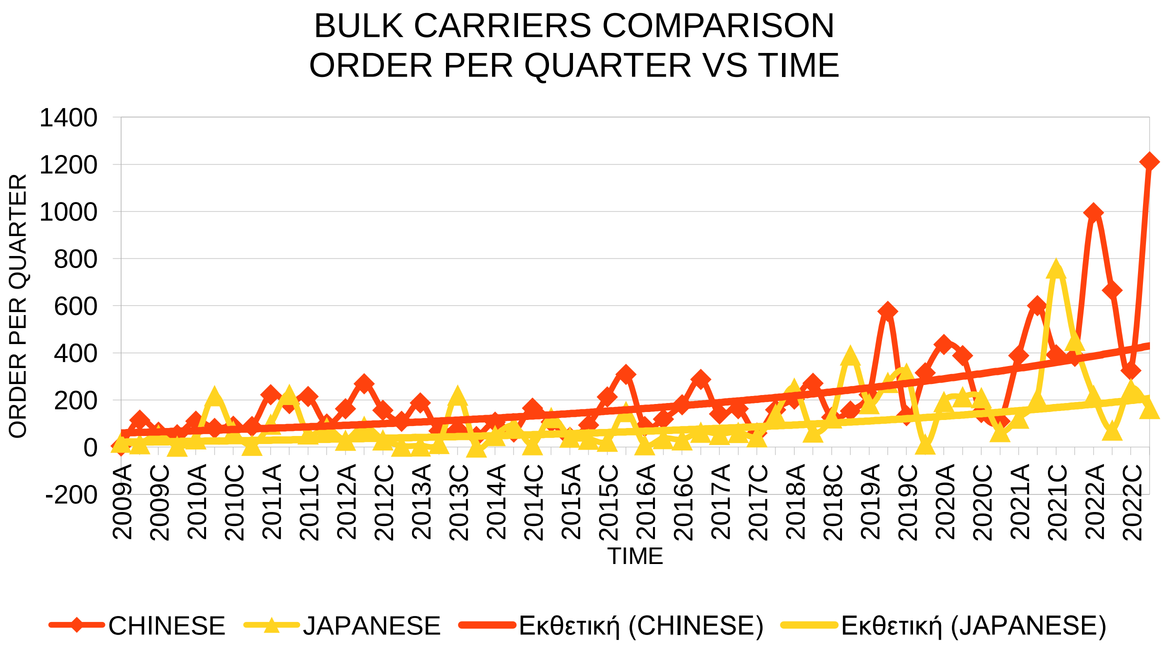
Plotting data only for the Chinese bulk carrier we get Figure 3, in which the trend line is present too. In Figure 3, by observing the data points and the trend line we see that as time passes the order value increases too, so there is an increasing trend which becomes sharper as time passes. There are several peaks and valleys but its seams that their occurrence is random and there is no seasonality/periodicity. There are several outliers that become bigger as time passes. Furthermore, there is no stationarity and data points present a big range. Last as far as variability is concerned, it is high since orders placed are not constant and high/low values exist through all quarters. We may assume that as time passes the Chinese vessel needs more spares and more maintenance to keep running smoothly. This may lead not only in increased expenses to buy the needed spares but to increased forwarding costs and man hours.

Figure 4. Japanese bulk carrier orders time plot



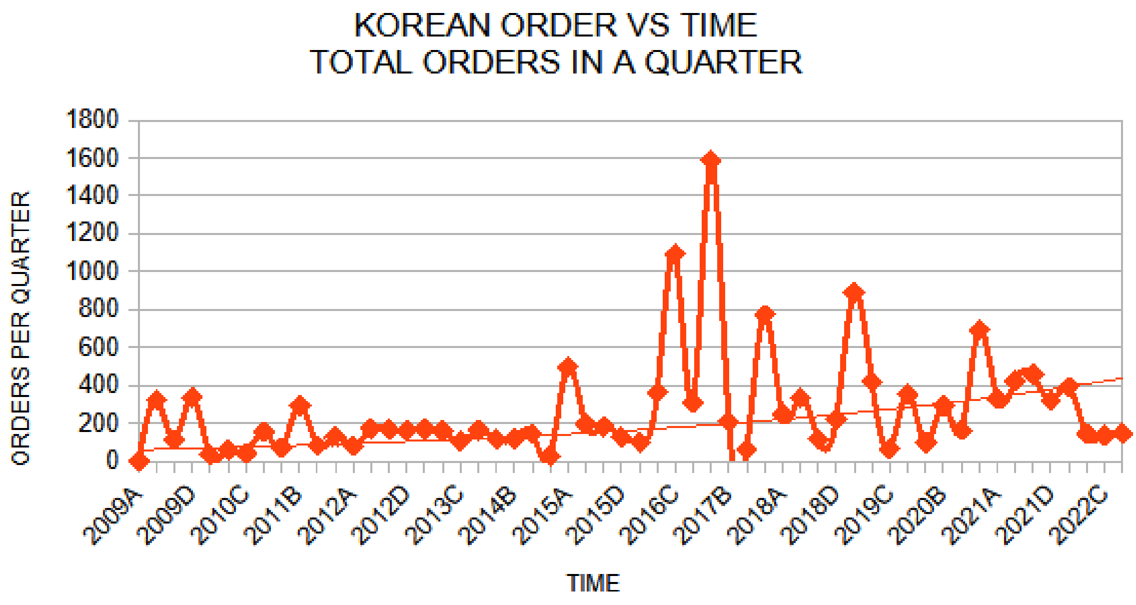
Same characteristics exist in Japanese bulker too, but in this case the trend inclination for the Japanese bulker is lower. Again, there are greater peaks as time passes but not at the same extend as happened for the Chinese bulk carrier. Another difference is that the range that data points belong too is smaller since the max value in Japanese bulk carrier is on about 750 and for the Chinese one is 1200. Seems that the expenses for the Japanese vessel are increased as time passes but not at the same extend as happens for the Chinese bulk carrier.

Figure 5. All bulk carriers' orders time plot



In above Figure 5, we may see a comparison for the two bulkers showing the inclination difference for the two origins which for the Chinese bulker is higher. At the final year, comparing the exponential, the expenses are more than the double for the Chinese compared to the Japanese. Based on the data retrieved, we may assume that this difference is even bigger since the average for the last year is almost five times bigger for the Chinese vessel compared to the Japanese one.

Figure 6. Korean tanker’s orders time plot



Last for the Korean tanker there is a increasing trend as seen in Figure 6, which is almost the same as was for the Chinese bulker. In this vessel, peaks occurred in the middle of the time, which may be translated as big machinery breakdowns which needed a lot of spares to make machinery operational or there might be a change in ordering policy that helped to tame the expenses.

In summarizing the insights derived from the aforementioned time plots, it becomes obvious that there does not exist any periodicity in the expenditure patterns across vessels of different origins and types. The irregular occurrence of peaks and valleys in the data implies a lack of recognizable cyclicality or systematic recurrence in spending behaviour.

Furthermore, the non-stationarity of the data indicates that the statistical properties of the expenditure series undergo changes over time, there is no stationarity. Notably, a visibly upward trend is observed in the cost associated with spare parts ordering, indicative of a consistent increase in expenditure. It is noteworthy that this ascending trend manifests distinctively among the three examined vessels even if there is similarity between the Chinese bulker and the Korean tanker.

The disparate trends in spare parts ordering costs for each vessel warrant a more thorough investigation, an examination that will be diligently undertaken in the following sections of this thesis. Trying to distinguish the factors that contribute the differing trends among the vessels is paramount for gaining comprehensive insights into the economic dynamics and operational considerations that impact expenses patterns for the technical maintenance of a ship.

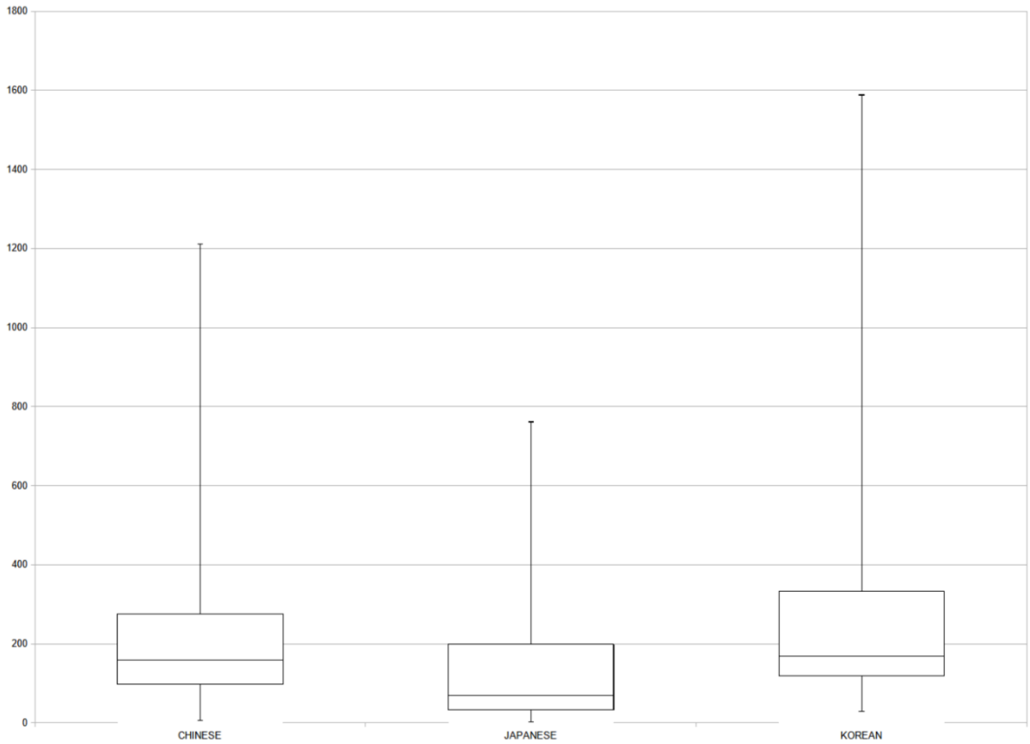
Next, we will create box plots for each vessel. In order to do so, we need to organize the data for the three different vessels and then calculate the minimum, the max, the median, the first and the third quartile. Table 2 below summarizes the results

Table 3. Min, Max and quartiles for all dataset

	CHINESE BULKER	**JAPANESE BULKER**	**KOREAN TANKER**
MIN	6.19	1.96	29.01
QUARTILE 1	98.16	32.27	118.36
MEDIAN	158.48	68.52	168
QUARTILE 3	276.19	198.28	332.78
MAX	1211.69	761.48	1588.61

A visual representation of the above values is shown below where a box plot for all three vessels is drawn.

Figure 7. All vessels box plot



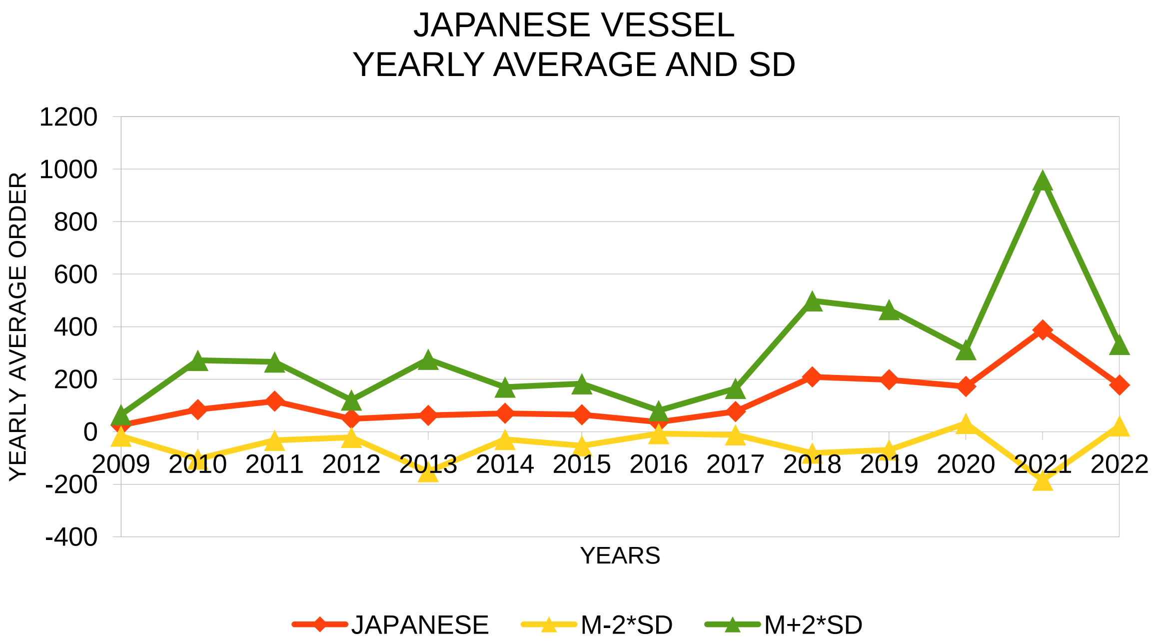
In the presented box plots, in Figure 7, significant distinctions emerge among the Chinese bulk carrier, Korean tanker, and Japanese bulk carrier concerning the distribution of order values. As far as the median is concerned, the Korean tanker demonstrates the highest median order value at $168, followed by the Chinese bulker at $158.48 and the Japanese bulker at $68.52. This suggests, on average, that for the Korean tanker incurs higher expenditure per order compared to the other two ships. Then for the interquartile range or IQR it is observed that the Korean tanker again exhibits the widest value for the interquartile range at 214.42 USD which means a bigger spread of orders' values. For the Chinese bulk carrier the IQR value is 178.03 USD which is less compared to the Korean tanker but again suggest the existence of some variability. Last for the Japanese bulker the IQR gets the lowest value equal to 166.01 USD which can be translated to a narrower range of operational needs, as far as spare parts is concerned and possible more consistent order sizes due to less breakdowns.

As far as the maximum and the minimum values observed in the boxplots, The Korean tanker has the highest maximum order value ($1588.61), suggesting it might have undertaken exceptional projects or faced specific high-cost maintenance needs due to excessive damages in its machinery. The Chinese bulker closely follows the Korean tanker with a maximum of $1211.69, indicating potential for significant individual orders. The Japanese bulker's maximum of $761.48 suggests its largest orders were considerably smaller compared to the other two ships. Lastly, regarding the presence of outliers, all three ships have outliers, highlighting instances where order values fell outside the typical range. The Korean tanker and the Chinese bulker have outliers reaching beyond $1000,

In summary, a first observation leads to the assumption that the Japanese bulk carrier tends to have lower-value orders. This inference is drawn from the fact that 50% of its orders are markedly lower in comparison to the Chinese bulk carrier and the Korean tanker. The Korean tanker exhibits the highest average and maximum order values, indicating potentially bigger maintenance needs. The Chinese bulker shows moderate variability with the potential for significant individual orders. The Japanese bulker demonstrates the most consistent ordering pattern with smaller average and maximum values. The box plots thus provide valuable insights into the comparative order value distributions among the three vessels, offering an understanding of their respective economic dynamics and potential operational implications.

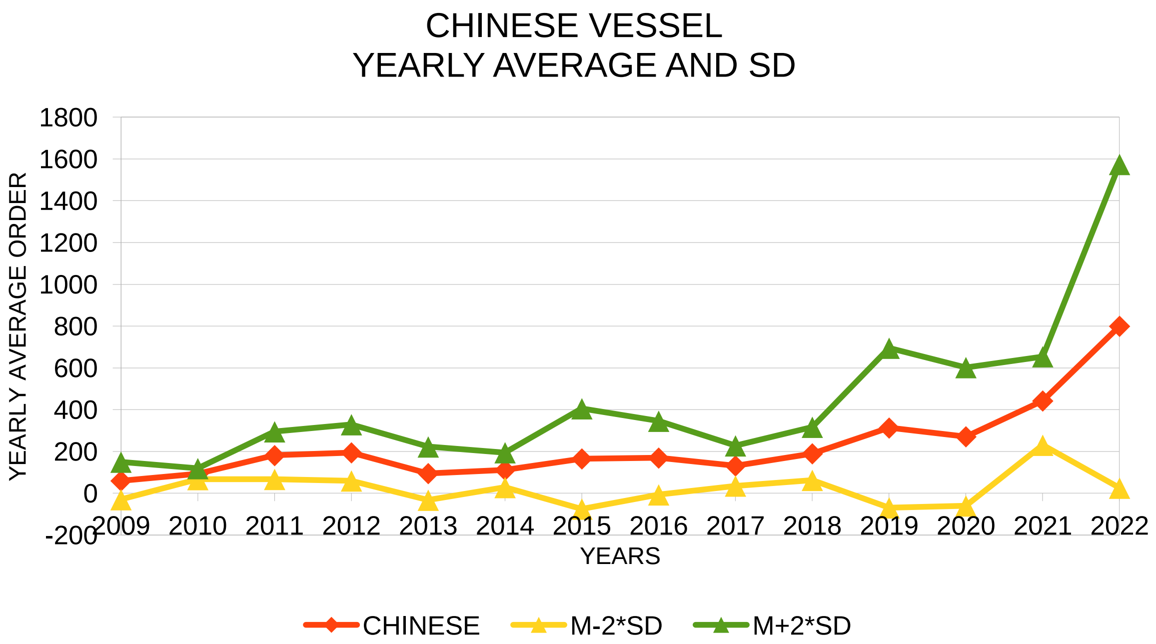
In the next plots we will assess the variability of the dataset over time by visualizing the average orders value for each year and the 2*S and -2*S where S stands for Standard deviation. Standard deviation is a measure used to find out how much does the data vary from the mean. A lower standard deviation implies that data points tend to be in proximity to the mean, indicating less variability. On the other hand, a higher value for the standard deviation means that data spread out over a broader range. The initial plot in diagram 9, focuses on the Japanese bulk carrier.

Figure 8. Japanese bulk carrier, yearly average and SD



Looking at the Japanese bulker data in Figure 8, we notice that both the average order size and the standard deviation increase over time, along with the distances covered. This suggests that the orders become more varied and spread out. However, after 2021, we see a decrease in both the average order size and the standard deviation. This shows a change from the previous trend, indicating that the orders may have become more consistent or stable during this time. These changes could be due to shifts in market conditions or how the ship is being operated, but more investigation is needed to understand why and more data after this time. Follows the plot for the Chinese bulk carrier in Figure 9,

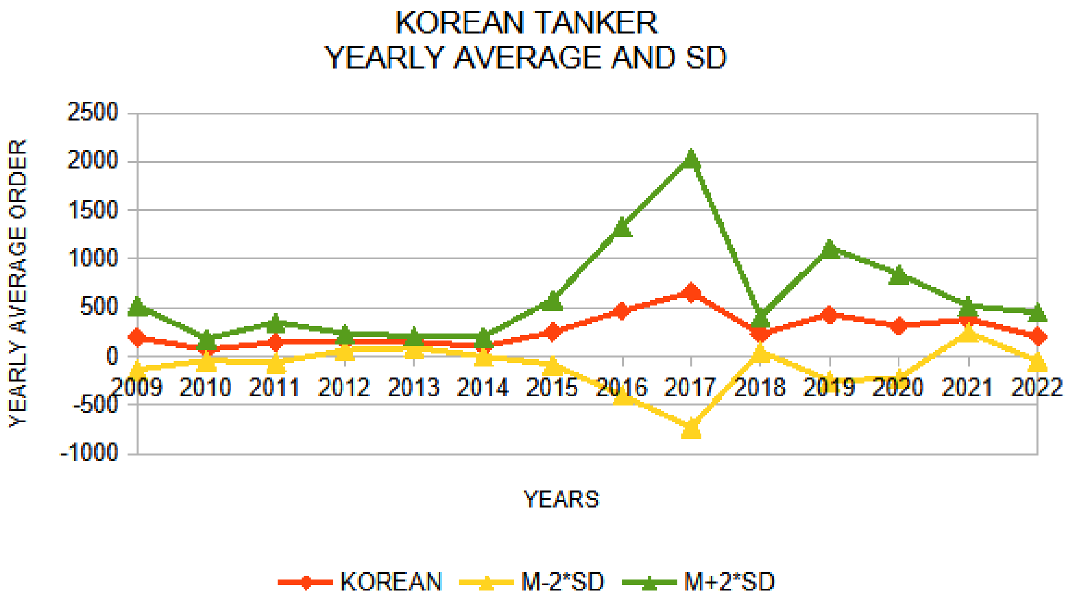
Figure 9. Chinese bulk carrier, yearly average, and SD



In Figure 9, we notice that the average order size over time increases more steeply compared to the Japanese bulker, and there isn't a point where it levels off or stabilizes. The same trend is observed for the standard deviation. This suggests that as time goes on, the orders become more varied and spread out compared to the Japanese bulker, as indicated by the increasing gap between the average and the standard deviation. When data points are more scattered like this, it can make predicting future expenses more challenging because there is less consistency or predictability in the ordering patterns.

In next plot the average for the Korean tanker is examined along with its deviation.

Figure 10. Korean tanker, yearly average, and SD



In Figure 10, we notice an outlier in the year 2017 on the graph. Up until that point, there is a rise in values followed by a sudden drop, particularly noticeable after 2019. It appears that both the average order size and the standard deviation stabilize after 2019, along with the distance covered. This suggests that there could have been a change in the company's suppliers' policy or perhaps a significant repair that improved conditions onboard, leading to more stable order sizes and less variability in the data in the following years.

6. Variables and model evaluation

Based on the insights gained from the preceding plots, it is evident that there is a significant correlation between the origin and type of a vessel and the corresponding order values. This suggests that factors such as the country of origin and the vessel type play a significant role in influencing the expenses incurred for orders. On the contrary, there appears to be a more limited association between quarters and order values, indicating that there isn't a clear pattern of periodicity in expenditure trends over time.

Moving forward, our analysis will proceed with deeper research into constructing regression models that encompass variables representing the vessel's origin, type, and quarters. These models will aim to quantify and evaluate the statistical relationships between each of these variables and the order values. By conducting these careful investigations, we seek to develop a comprehensive understanding of how factors such as the origin, type, and age of a vessel interact with the order values.

Our analysis is going to use the regression model analysis which is a statistical method used to understand the relationships between different variables. By examining how one variable affects another, regression analysis helps to uncover patterns, trends, and associations within data. This analytical approach is widely utilized across various fields, including economics, social sciences, healthcare, and engineering, among others. By quantifying the relationships between variables, regression analysis provides valuable insights into the underlying mechanisms driving observed phenomena. This helps researchers, analysts, and decision-makers make informed predictions, identify influential factors, and assess the impact of interventions or changes. Overall, regression model analysis serves as a versatile and powerful tool for uncovering insights, making predictions, and informing decision-making processes across diverse contexts.

Before proceeding with the regression analysis outlined in the final model, we conducted a thorough examination of other potential models. This exploration commenced with an initial examination of a model containing solely a linear variable. This linear model serves as a foundational starting point for our analysis, allowing us to establish a baseline understanding of the relationship between the variables of interest and the order values. Subsequent repetition of the model will incorporate additional variables and complexities to further refine our understanding and capture the nuanced interplay between the various factors influencing order values within the maritime industry.

6.1 Simple linear regression analysis

In this model, we examine the possibility of the dependent variable $\hat{y}$, which in our case is the order value in a year's quarter, being dependent on the time that has passed or else the age of the vessel. The general form of this model is

$$\hat{y} = b_0 + b_1 x + e$$

where x is each year's quarter and e is the error term. Running the regression model analysis for this case we get the results shown in table 4

Table 4. Simple Regression model analysis

	Chinese bulker		Japanese bulker		Korean tanker	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Intercept	4.283	5.89e-22	3.20	1.65e-10	4.64	8.27e-17
x	0.023	0.017	0.029	0.068	0.026	0.059

If the p-value associated with a variable is below your chosen significance level, it indicates that the sample data present sufficient evidence to reject the null hypothesis for the entire population (Frost, 2023). In our case the significance level is equal to 5%, The data support the hypothesis that there is a correlation that is not equal to zero. At the population level, alterations in the independent variable are linked to changes in the dependent variable. This particular variable holds statistical significance and is likely a valuable inclusion in your regression model.

In our regression analysis we see that both p-values are below 0.05 only for the Chinese bulker, indicating correlation between time passing or age of ship and order value. For the other two vessels the P-value for the coefficient for the x, which is the passing year's quarters or ship's age, is marginally larger the 0.05 and so we assume that there is no statistically connection between passing time and order value for the Japanese bulker and the Korean tanker.

6.2 Multiple linear regression for origin of vessel

Next, we will examine a model in which the independent variables shall represent the age of the ship and its origin. For this reason, except of the linear variable we will have one dummy variable $CHINESE$ and the data that will be used will be only for the two bulk carriers in order not to have vessel of another type. When the data refers to the Chinese bulk carrier the variable $CHINESE$ will be equal to 1 and for the rest cases will be equal to zero. The general format of the model shall be

$$\hat{y} = b_0 + b_1x + b_2CHINESE + e$$

Regression analysis for the above model gives the results presented in table 5

Table 5. Multiple linear regression for origin of vessel results

	Coefficient	P-value
Intercept	3.0932	5.99e-31
Trend	0.0385	2.89e-11
Chinese	0.8898	6.32e-07

In these results, we see that all the P-values are less than the significance level of 5% so we accept the hypothesis that there is statistical connection between dependent variable of order value and independent variables of time and origin and so variables will be kept if our model and its final form will be

$$\hat{y} = 3.0932 + 0.0385x + 0.8898CHINESE$$

6.3 Multiple linear regression for type of vessel

Following this, we will explore a model where the independent variables denote the age of the ship and its type. In a similar way to previous model, a single dummy variable, $TANKER$, will be incorporated. The dataset employed for this analysis will involve all three vessels. When the data pertains to the Korean tanker, the $TANKER$ variable will assume a value of 1, while for all other instances, it will be set to zero. The overall structure of the model will be as follows:

$$\hat{y} = b_0 + b_1x + b_2TANKER + e$$

Regression analysis for the above model gives the results presented in table 6

Table 6. Multiple linear regression for type of vessel results

	Coefficient	P-value
Intercept	3.7542	4.63e-54
Linear	0.0310	3.75e-10
Tanker	0.6773	3.50e-05

The results show that all the P-values are lower than the 5% significance level. This means we accept the hypothesis that there is a statistical connection between the dependent variable (order value) and the independent variables (time and type). Therefore, these variables will be retained in our model, and they will be part of its final form.

$$\hat{y} = 3.7542 + 0.031x + 0.6773\, TANKER$$

6.4 Multiple linear regression for origin, type of vessel and quarters

In this particular model, we aim to investigate the relationships among origin, type, and time. To capture the temporal aspect, three dummy variables—Q2, Q3, and Q4—are used. Specifically, Q2 takes on a value of 1 when the order pertains to the second quarter of a year, and similarly, Q3 and Q4 operate under the same principle for the third and fourth quarters, respectively. All three quarters assume a value of zero when the order is placed in the first quarter. For the first quarter we do not use a separate dummy variable to avoid the "dummy variable trap," (Mohanan, 2023) which occurs due to perfect multi-collinearity when all the dummy variables are included, leading to issues in the estimation of the regression coefficients. This use of dummy variables allows us to examine the periodicity within the dataset, checking for any correlation of one or more quarter to the independent value.

For the type and origin, we employ the variables $CHINESE$ and $KOREAN$. These variables cover all possible scenarios: if the vessel is a bulker and Chinese, the $CHINESE$ variable is set to 1; if it is a tanker, the $KOREAN$ variable takes on a value of 1, and if the vessel is a Japanese bulker, both variables are set to zero, so a third dummy is not needed again to avoid the “dummy variable trap”.

Additionally, we incorporate a linear connection variable to examine linear relationships within the dataset. Running the regression model analysis based on the described model leads to results presented in Table 7.

Table 7. Regression model coefficients and P-values

	Coefficients	**P-value**
Intercept	3.286	1.002 E-34
Trend	0.031	2.037 E-11
Chinese	0.890	5.879 E-07
Korean	1.122	6.902 E-10
Q2	0.190	0.337
Q3	-0.002	0.99
Q4	-0.124	0.531

In regression analysis for each variable there are two hypotheses, the null-hypothesis that there is no relationship between the examined variable and the dependent variable, in which in our case is the orders value and the alternative hypothesis that there is relationship. To accept the null hypothesis the P-value for the variable must be above 0.05 is a significance level of 95%. The significance level of 95% is the most used. In above table, we see that the P-value for all quarters' variables, Q2, Q3 and Q4, is bigger than 0.05 so we accept the null hypothesis, which aligns with what we anticipated based on the time plots examined in Chapter 6. For the rest variables, the P-value is lower than the 0.05 so there is statistically significant link between the variables and the dependent variable. Since the variables for the time do not have significant statistical relationship with the dependent variable, we may say that there is no indication of periodicity for the order value. In that case the model will be the below

$$\hat{y} = 3.2857 + 0.0312t + 0.8898CHN\_BLK + 1.1222KRN\_TNK$$

in which t is the equal to 1 for first quarter of 2009, 2 for the second of 2009 and continues in the same pattern, CHN_BLK takes the values 0 and 1, 1 when the vessel is Chinese bulk carrier and zero for the rest cases. For the variable KRN_TNK is like CHN_BLK with the difference that becomes equal to 1 when the vessel is a Korean tanker.

To evaluate model forecasting ability, we will have to use it in the forecasting portion of our dataset which is going to be the last 18 values for each vessel, so it is the third quarter of year 2018 until the fourth quarter of year 2022. By using the model we get the estimated value of the orders value for this quarter and then by comparing to the observed value we calculate the two indicators MAE (Mean Absolute Error) which is the average of the absolute difference between the predicted and the actual value and the MAPE (Mean Absolute Percentage Error) which is the average of the absolute percentage of the difference between the predicted value and the actual value. The results we get are shown in the table 4 below;

Table 8. MAE and MAPE values

MAE	**MAPE**
0.64	12.41

As per Lewis (1982) a value for MAPE between 10 and 20 indicates a good forecasting ability for a model. Same we may say for MAE since the mean values for the order is 5.43 and MAE is 0.64.

7. Conclusions and future research opportunities

7.1 Conclusions

The following four research questions were initially set;

1. How much are expenses related to type of vessel
2. How much are expenses related to origin of vessel
3. Is there any periodicity/seasonality observed in the data? If yes at what period?
4. What is the accuracy of a forecasting model?

As far as the first two questions is concerned, regression model analysis and time plots of expenses versus time, revealed a robust and statistically significant correlation between expenses and vessel characteristics of type and origin. Both methods demonstrated that the Chinese bulk carrier as time passes has a steeper increase in expenses compared to the Japanese bulk carrier and the Korean tanker. Moreover, the data for the Chinese bulk carrier as time passes become more scattered. Such knowledge may assist in optimizing financial strategies, operational efficiency, and investment payback.

Regarding the third research question of periodicity/seasonality of data, thorough examination of the dataset revealed absence of the same. Neither time plots nor the regression model analysis showed any periodicity or seasonality for the expenses, which means that there is a consistent increase in expenses that may stem from factors such as aging equipment and increased usage rather than being influenced by seasonal factors such as week, month or year.

Lastly, regarding the accuracy of the proposed model, by using metrics such as Mean Absolute Percentage Error (MAPE) and Mean Absolute Error (MAE) has been evaluated as good. Using the regression model we may have good forecasting result based on the characteristics of the vessel (origin and type).

7.2 Future research opportunities

While the proposed models exhibit a good accuracy, there is a potential for further refinement. Future research could use a more extensive dataset that may include more details such as origin of spares for example, and in that way find a better coefficient for each variable or increase the variables used and in that way the forecasting ability of the model. An interesting future research direction would be to apply a nonlinear regression model with adequate flexibility in the functional forms it can accommodate. An obvious candidate is the artificial neural network (ANN). Although these devices can in principle detect highly nonlinear relationships among data variables, they are prone to overfitting. Carefully selecting the optimal design of the ANN is of paramount importance for its robustness and good performance out-of-sample. Thomaidis and Dounias (2011, 2012) discuss various approaches to ANN model selection based on sequential statistical tests.

An aspect of our study that can be investigated in the future is the time point that a vessel becomes not profitable and thus should be sent for scrap. This point should be different for each vessel based on each vessel's characteristics. Since a vessel is a group of machineries on a ship, we may compare the expenses needed for the maintenance of a vessel with those of another industry and try to figure out similarities that will help in financial strategies followed.

References

Admin. (2022, April 21). *Measures of Dispersion in Statistics (Definition & Types)*. BYJUS. https://byjus.com/maths/dispersion/

Egyptian Ships - Ages of Exploration. (2014, October 2). Ages of Exploration. https://exploration.marinersmuseum.org/watercraft/egyptian-ships/

Frost, J. (2023, March 21). *How to interpret P-values and coefficients in regression Analysis*. Statistics by Jim. https://statisticsbyjim.com/regression/interpret-coefficients-p-values-regression/

Gershon, L. (2021, April 28). Is This 10,000-Year-Old Carving Europe's Oldest Known Depiction of a Boat? *Smithsonian Magazine*. https://www.smithsonianmag.com/smart-news/10000-year-old-boat-carving-found-norway-180977611/

Lewis C. D. (1982a). *Industrial and business forecasting methods : a practical guide to exponential smoothing and curve fitting*. CiNii Books. http://ci.nii.ac.jp/ncid/BA12260589

Mohanan, M. (2023). Understanding the dummy variable trap in regression models. *ResearchGate*. https://doi.org/10.13140/RG.2.2.17430.52806

Plomaritou, E., &Jeropoulos, S. (2022). The digitalisation in chartering business: special reference to the role of e-bill of lading in the bulk and liner markets. *Journal of Shipping and Trade*, *7*(1). https://doi.org/10.1186/s41072-022-00129-2

Sethuraman, M. (2011). Measures of central tendency: The mean. *Journal of Pharmacology and Pharmacotherapeutics*, *2*(2), 140–142. https://doi.org/10.4103/0976-500x.81920

Ship | Definition, Types, Old, & Facts. (1998, September 8). Encyclopedia Britannica. https://www.britannica.com/technology/ship/Types-of-ships

Steamboat. (n.d.). https://education.nationalgeographic.org/resource/steamboat/

Wikipedia contributors. (2024, January 27). *Junk (ship)*. Wikipedia. https://en.wikipedia.org/wiki/Junk_(ship)

Thomaidis, N.S., &G. Dounias (2011), "On detecting the optimal structure of a neural network model under strong statistical features in errors", Journal of Time Series Analysis 32 (3), pp. 204-222.

Thomaidis, N.S., &G. Dounias (2012), "A comparison of statistical tests for the adequacy of a neural network regression model", Quantitative Finance 12(3), pp. 437-449.